

Leveraging Gait for People Perception

by Julien Claude Guy Collart

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

under the supervision of Dr. Alen Alempijevic and Dr.
Teresa Vidal-Calleja

University of Technology Sydney
Faculty of Engineering and Information Technology

January 2020

Declaration of Authorship

I, Julien Claude Guy Collart, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:

Signed: Signature removed prior to publication.

Date: 21/01/2020

UNIVERSITY OF TECHNOLOGY SYDNEY

Abstract

Faculty of Engineering and Information Technology

Centre for Autonomous Systems

Doctor of Philosophy

by Julien Claude Guy Collart

In the past few years, there has been an increase in the number of robots interacting with people in public spaces. Robots and humans co-existing in an environment require an understanding of each other's intention to perform safe and optimal interactions. One of the most basic forms of intention inference consists of predicting an individual's future poses. While existing work on human tracking tends to use basic knowledge of the gait directly extracted from the lower-body that governs the spatio-temporal dynamics of the locomotion, such methods become impractical in public environments where the number of people and occlusions make it nearly impossible to see the placement of the feet and legs.

In this work, we propose a framework for people perception and intention inference in real-time which leverages the relationship between the feet and upper-body for estimating and predicting future postures and gait phases when the feet cannot be directly observed. In addition, the inputs to the framework solely use depth data, with the added benefit of being privacy-friendly.

The spatio-temporal relationship between the feet and the upper-body was exploited to design a framework for gait phase inference from shoulder pose observations. Classification of gait phases is achieved using Hierarchical Hidden Markov Models that are capable of modelling complex sequences. The experimental results confirm that the current gait phase can be estimated from shoulder poses alone.

Given that shoulder motion can be used for inferring the current gait phase, a second framework was designed to exploit the anticipatory signals contained in the upper-body motion to predict the future gait phase and posture. A Recurrent Neural Network was used to encode long-term dependencies in the sequences of poses. A custom loss function was constructed to direct the learning by incorporating the bio-mechanical constraints of human locomotion. Experimental results confirm that the motion states, postures and gait phases, can be predicted from a sequence of shoulder poses, and that the accuracy is improved by the custom loss function.

Finally, the problem of predicting an individual's future poses is addressed by transferring the learning from the second framework's model to train another classifier used to drive a dynamic motion model's internal parameters within a tracking algorithm. Experimental results confirm that the motion states predictive model can be used to boost and improve the learning of a different classifier related to human motion, so that human tracking is improved.

Acknowledgements

I want to take this opportunity to thank so many people who contributed to helping me while probably not even realizing.

First, I would like to thank my supervisors, Dr. Alen Alempijevic and Dr. Teresa Vidal-Calleja, for all their support and guidance along this long and challenging journey. Their motivation and expertise allowed me to become who I am now. I would also like to thank Dr. Nathan Kirchner, who made me come to Australia and guided me through the initial steps of my Ph.D.

I would like to thank the Rail Manufacturing CRC for funding my research. It has been a great pleasure working with you on this project. You always cared for your students. I would like to particularly thank the director, Dr. Stuart Thomson, and Katie Rizo for everything. As well as all the students I have met during the workshops.

日本でのインターンシップを一年間サポートしてくれた株式会社ホンダ・リサーチ・インスティテュート・ジャパン、特にランディ・ゴメス博士に感謝します。そこで出会った友人たち、しおのさん、ランディ・ジャクソン、マーク・サントス、そしてインターンの皆さんに感謝します。

There are many people met in Australia that I want to thank. People from UTS-CAS: academics and students who taught me a lot about research and life; And some of my friends who had to deal with my gigantic craziness: Phillip Quin, Lakshitha Dantanarayana, Katherine Waldron, Tamsin Claire, Richardo Khonasty, Maleen Jayasuriya, Kyle Alvarez, Marija Popović, Stefano Aldini, Alex Virgona, Stefan Kiss, Karthick Thiyagarajan, Sheila Sutjipto, James Poon, Robert Yang, and Jason Schenkenberg.

Je voudrais aussi remercier des gens, qui même à des milliers de kilomètres, ont su rester proche de moi : ma famille et mes amis. Je sais que je n'ai pas dû être facile à supporter. Faire une thèse s'est révélé bien plus difficile que je ne l'avais envisagé. Et pour cette raison, je voudrais remercier certaines personnes en particulier qui ont su m'épauler dans les moments durs : Léo Ghafari, Camille De Carvalho, les Boczulak (Julien, Vero, Bobo), les Mailhe (Cyrille, Odile, Dorian, Lisa), Nicolas Autier, Mark Thebault, Coline Delefolly, Hélène Monchaux, Mélanie Romero, et Michael Pérez.

Enfin, je voudrais adresser un remerciement très spécial à ma soeur, Sandie Collart. Merci pour tout ton soutien sans faille dans cette aventure.

En mémoire de ma mère. Tu me manques..

Catherine Collart, née Clement

(2 Août 1962 - 31 Août 2012)

Contents

Declaration of Authorship	iii
Abstract	iv
Acknowledgements	vi
List of Figures	xiii
List of Tables	xv
Glossary of Terms	xx
1 Introduction	1
1.1 Research Problem	1
1.2 Background and Motivation	2
1.3 Research Questions	6
1.4 Contributions	6
1.5 Publications	7
1.6 Structure of the Dissertation	7
2 Background	9
2.1 Introduction	9
2.2 Human Locomotion	9
2.2.1 Quiet Standing	10
2.2.2 Gait Initiation	11
2.2.3 Gait Cycle	12
2.2.4 Gait Termination	13
2.2.5 Turning	14
2.3 Gait Analysis	17
2.4 Summary	19
3 Gait Phases Estimation using Graphical Models	21
3.1 Introduction	21
3.2 Related Work	22

3.3	Background: Graphical Models	24
3.3.1	Multivariate Gaussian Mixture Models	24
3.3.2	Hidden Markov Models	25
3.3.3	Hierarchical Hidden Markov Models	27
3.4	Gait Phases inference using Hierarchical Hidden Markov Models with Multivariate Gaussian Mixture Outputs	30
3.4.1	Feature Extraction	31
3.4.1.1	Offline Extraction	33
3.4.1.2	Online Extraction	34
3.4.2	Hidden Markov Models with Mixture of Gaussian Outputs for Sub-States Inference	35
3.4.2.1	Gait Cycle	36
3.4.2.2	Gait Initiation	37
3.4.2.3	Gait Termination	40
3.4.3	Hierarchical Hidden Markov Models for Motion States Inference	41
3.5	Experimental Results	42
3.5.1	Hidden Markov Models Evaluation	45
3.5.2	Hierarchical Hidden Markov Models Evaluation	47
3.6	Limitations	49
3.7	Summary	50
4	Motion States Prediction using Recurrent Neural Networks	51
4.1	Introduction	51
4.2	Related Work	53
4.3	Background: Recurrent Neural Networks	54
4.3.1	Feed-Forward Neural Network	54
4.3.2	Long Short-Term Memory Network	56
4.3.3	Mini-batch Gradient Descent Optimizer	59
4.4	Modeling and Predicting the Motion States using Long Short-Term Memory Networks	60
4.4.1	Feature Extraction	60
4.4.2	Motion States	61
4.4.3	Model Architecture	62
4.4.4	Custom Loss Function	63
4.5	Experimental Results	64
4.5.1	Experimental Parameters	65
4.5.1.1	Data Preparation	65
4.5.1.2	Network Training Properties	67
4.5.2	Custom Loss Evaluation	67
4.5.3	Prediction Model Evaluation Without Noise	69
4.5.4	Prediction Model Evaluation With Noise	74
4.6	Discussion	76
4.7	Summary	77
5	Application to Human Tracking	79

5.1	Introduction	79
5.2	Related Work	80
5.3	Leveraging the Motion States Prediction Model for Predicting Changes in the Dynamics	82
5.3.1	Velocity States Definition	82
5.3.2	Transfer Learning for Velocity States Prediction	86
5.3.3	Experimental Evaluation	86
5.3.3.1	Transfer Learning versus Regular Learning	87
5.3.3.2	Velocity States versus Motion States	89
5.4	Improving Human Tracking with Velocity States	90
5.4.1	Human Tracking	90
5.4.1.1	Detection	90
5.4.1.2	Filtering	92
5.4.1.3	Evaluation Methods	94
5.4.2	Experimental Evaluation	94
5.4.2.1	Evaluation on Motion Capture Data	95
5.4.2.2	Evaluation on Point Cloud Data	99
5.4.2.3	Application on Train Station Data	103
5.5	Summary	109
6	Conclusion	111
6.1	Summary of the Contributions	111
6.2	Discussion of Limitations	113
6.2.1	Features Standardizing From A Cycle	113
6.2.2	Limited Features	113
6.2.3	Adapting the Motion Model's Parameters	113
6.3	Future Research	114
6.3.1	Person Re-identification	114
6.3.2	Velocity Prediction using Recurrent Neural Networks	114
6.3.3	Exploring Additional Features and Usages	115
	Appendices	116
A	Experimental Data Acquisition	117
A.1	Motion Capture Environment	117
A.2	Markers Placement	118
A.3	Experimental Scenarios	118
A.3.1	Straight Line Walk	119
A.3.2	Multi-directional Walk	120
A.3.3	Social Walk 1	121
A.3.4	Social Walk 2	121
A.4	Experimental Datasets	122
A.4.1	Dataset A	124

A.4.2	Dataset B	124
A.4.3	Dataset C	125
A.4.4	Dataset D	126
A.4.5	Dataset E	126
A.4.6	Dataset F	127

Bibliography

List of Figures

1.1	Non-interference headway used for safe separation between trains	3
1.2	Example of dwell-time management by a rail employee required for a single door.	4
1.3	Example of people represented as point cloud.	5
2.1	State machine of the main actions performed during human walking.	10
2.2	Gait cycle phases.	13
2.3	Turning strategies while walking.	16
3.1	Example of Hierarchical Hidden Markov Models model illustrated as a tree	27
3.2	Example of Hierarchical Dynamic Bayesian Network	28
3.3	Motion States dependencies.	31
3.4	Axis definition in the world frame and the person frame.	32
3.5	Four people GCY offline features before standardizing.	33
3.6	Four people GCY offline features after standardizing.	34
3.7	Four people GCY online features after standardizing.	35
3.8	Gait Cycle sub-states.	36
3.9	Feet measurements for GCY after alignment.	37
3.10	Offline features for GCY after alignment.	38
3.11	Online features for GCY after alignment.	38
3.12	Gait Initiation sub-states.	39
3.13	Feet measurements for GIL after alignment.	39
3.14	Offline features for GIL after alignment.	40
3.15	Online features for GIL after alignment.	41
3.16	Gait Termination sub-states.	41
3.17	Feet measurements for GTR after alignment.	42
3.18	Offline features for GTR after alignment.	43
3.19	Online features for GTR after alignment.	44
3.20	HHMM tree structure.	44
4.1	Multi-Layer Perceptron.	55
4.2	An example of an unrolled Recurrent Neural Network.	57
4.3	Types of Recurrent Neural Networks.	57
4.4	Long Short-Term Memory Cell.	58
4.5	Features extraction for motion states prediction.	61
4.6	Example of unrolling for a Many-To-One Recurrent Neural Network.	62

4.7	Motion States prediction network architecture.	63
4.8	Transitions matrix between the motion states.	65
4.9	Example of labeling for straight motion.	66
4.10	Example of labeling for turning to the left while walking.	67
4.11	States transitions improvement using the custom loss function.	69
4.12	Motion states prediction network training without noise.	70
4.13	F1-score per states for motion states prediction without noise.	72
4.14	Confusion matrix for motion states prediction without noise.	73
4.15	Motion states prediction network training with noise.	74
4.16	F1-score per states for motion states prediction with noise.	75
4.17	Confusion matrix for motion states prediction with noise.	76
5.1	Example of shoulders 3D magnitudes labeling.	83
5.2	Example of empirical thresholds extraction for the Velocity States.	84
5.3	Example of filtering with Savitzky-Golay on the angular velocity.	85
5.4	Network architecture for velocity states prediction.	86
5.5	Improvement of the confusion matrix for velocity states prediction without noise.	89
5.6	Evaluation of the learning of velocity states classifier.	90
5.7	Example of ellipsoids fitting on point-cloud data.	91
5.8	Example of improvement in the tracking with ACCE-V prediction.	98
5.9	Example of improvement in the tracking with TR-V prediction.	99
5.10	Example of improvement in the tracking with TRS-V prediction.	100
5.11	Example of improvement in the tracking with DECE-V prediction.	101
5.12	Example of tracking failure due to incorrect prediction.	102
5.13	Camera placement on a platform at Brisbane Rail Station.	104
5.14	Example of tracking in a train station: a person walking through.	104
5.15	Example of tracking in a train station: a person stopping.	105
5.16	Example of tracking in a train station: a person stopped.	105
5.17	Example of a passenger's abrupt turn on a platform at Brisbane Rail Station.	106
5.18	Example of improvement with our method on a real-world train station dataset.	107
5.19	Comparison of the number of tracks counted on a real-world train station dataset.	108
A.1	UTS Data Arena used for passive motion capture.	118
A.2	IR-reflective markers attached on a participant.	119
A.3	Example of the dataset collected from a participant following the straight line walk scenario.	120
A.4	Example of the dataset collected from a participant following the multi- directional scenario.	121
A.5	Example of the dataset collected using the social walk 1 scenario.	122

List of Tables

3.1	States and Sub-States for gait inference.	22
3.2	Manual labeling of the experimental dataset using the feet measurements per states/sub-states.	45
3.3	Accuracy for Offline inference using the feet features per model.	46
3.4	Accuracy for Offline inference using the shoulders features per model. . . .	47
3.5	Accuracy for Online inference using the shoulders features per model. . . .	48
3.6	Accuracy for Online states inference using the shoulders features with full HHMM model.	48
4.1	List of the Motion States predicted.	52
4.2	F1-score average computed on data without noise.	72
4.3	F1-score average computed on data with noise.	75
5.1	Velocity States used in this chapter.	80
5.2	Thresholds for labelling the Velocity States using the shoulders' linear ve- locities and accelerations.	84
5.3	Thresholds for labelling the Velocity States using the shoulders' angular velocities.	85
5.4	Velocity states prediction with noise.	88
5.5	Parameters used for the Level 1.	96
5.6	Parameters used for the Level 2.	96
5.7	Parameters used for the Level 3.	96
5.8	Tracker performance for 3 evaluation methods with noise on <i>DATASET_D</i> . . .	97
5.9	Tracker performance for 3 evaluation methods with noise on <i>DATASET_F</i> . . .	97
5.10	Tracker performance for 3 evaluation methods with noise on <i>DATASET_E</i> . . .	98
5.11	Parameters used for the experimental evaluation of the point cloud (Level 1). .	100
5.12	Parameters used for the experimental evaluation of the point cloud (Level 2). .	101
5.13	Tracker performance evaluated on <i>DATASET_E</i> Kinect data.	103
5.14	Tracker performance evaluated on <i>DATASET_F</i> Kinect data.	103
A.1	Markers and scenarios recorded per participant.	123
A.2	Experimental Dataset A (<i>DATASET_A</i>).	124
A.3	Experimental Dataset B (<i>DATASET_B</i>).	125
A.4	Experimental Dataset C (<i>DATASET_C</i>).	126
A.5	Experimental Dataset D (<i>DATASET_D</i>).	126
A.6	Experimental Dataset E (<i>DATASET_E</i>).	127

A.7 Experimental Dataset F (*DATASET_F*). 127

Acronyms & Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
CAS	Centre for Autonomous Systems
UTS	University of Technology Sydney
COM	Center Of Mass
COP	Center of Pressure
MAXCOP	Maximum Center Of Pressure
GI	Gait Initiation
GCY	Gait Cycle
GT	Gait Termination
GIR	Gait Initiation with Right foot first
GIL	Gait Initiation with Left foot first
GTR	Gait Termination with Right foot first
GTL	Gait Termination with Left foot first
GMM	Gaussian Mixture Model
EM	Expectation-Maximization

E-Step	Expectation Step
M-Step	Maximization Step
HMM	Hidden Markov Model
Pr	Probability
HHMM	Hierarchical Hidden Markov Models
H-DBN	Hierarchical Dynamic Bayesian Network
RH	Right foot Heel
LH	Left foot Heel
RT	Right foot Toes
LT	Left foot Toes
LOO	Leave One Out
STA	Standing
TRS	Turning on the Spot
TRSR	Turning on the Spot to the Right
TRSL	Turning on the Spot to the Left
TRR	Turning Right while walking
TRL	Turning Left while walking
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory Network
SGD	Stochastic Gradient Descent
TP	True Positive

TN	True Negative
FP	False Positive
FN	False Negative
RMSE	Root Mean Square Error
MOTP	Multiple Objects Tracking Precision
MOTA	Multiple Objects Tracking Accuracy
IR	Infrared
FPS	Frames Per Second
HE	Head
SH	Shoulders
STR	Straight walk scenario
MDI	Multi-direction walk scenario
SO1	Social walk scenario 1
SO2	Social walk scenario 2
ID	Identity

Glossary of Terms

Autonomous	Without human intervention.
Shoulders	Trapezius muscle.
Pose	Position and orientation.
Postures	Non-forward and non-backward motion.
Gait Phases	Phases shown during locomotion.
Motion States	States describing either only the gait phases or both.
Sub-state	Internal state to a state.

